

# Outline Technical Framework For Bastille Dérive and Other Urban Fictions

Atau Tanaka and Petra Gemeinboeck

There will be two types of participants in Bastille Dérive, corresponding to the two spaces the work inhabits: the outdoor urban environment and the indoor gallery environment. Mobile players (outdoors) will be armed with advanced mobile phones loaned out to participants. The gallery participants will not only experience the resulting audio/visual projections in the gallery, but will interact with the piece and the other participants through touch-screen interfaces mounted on pedestals in the gallery space.

## Data Aggregation

The system server will aggregate the different audio video sources from the mobile devices, and will process them according to control information coming from the different participants and various external sources. Mobile participants will be able to interact in a crude fashion using the 5-way control buttons on their mobile phone. In a more advanced implementation, sensors onboard the mobile devices will provide more fluid information on the gesture and movement of the mobile user. While these various control data will directly affect the audio/visual

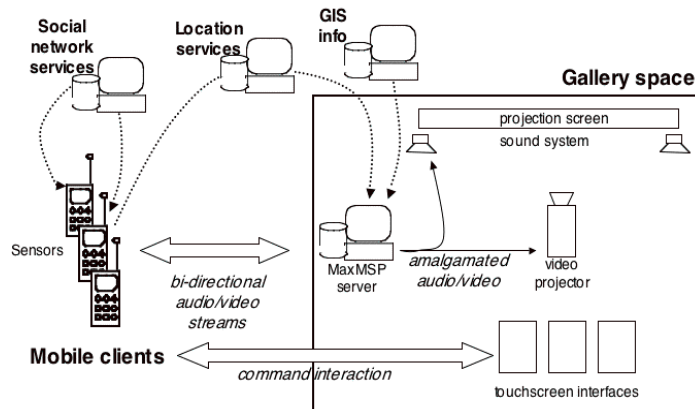


Figure. System Overview

display local to their device, their primary use will be to be recovered by the system as a series of XML messages that alter that participant's part within the aggregated whole on the server.

Likewise, gallery participants actions on the touch-screens will not only serve to send instructions to their corresponding mobile user, but will also be sent up to the server to contribute to global media processing. External control sources include geographic location of each mobile device. While existing systems like the Global Positioning System (GPS) allow accurate location tracking, they necessitate additional hardware. In the current framework, we make use of mobile phone's GSM cell antenna signal reception strength mapped to mobile operators' databases of antenna location to deduce an approximate location. A Location Based Services (LBS) server provides the gateway between the measured antenna signal, the operator database, and the framework server. Other external information (deployed in <Impossible Geographies: Urban Fiction>) will include Geographic Information Systems (GIS). This will allow recovering demographic information correlated to mobile client position.

## System Output

Based on aggregated media and associated control information, the system server will amalgamate the disparate media elements, processing, combining, mixing, distorting them according to rules and mappings defined in each specific embodiment. The result will be a series of audio/video streams sent out to several destinations. The destinations include returns to the individual mobile devices and in-gallery touch screen displays. The primary destination is the main video and sound projection in the gallery, representing the sum output of the system. The media being streamed out to the mobile devices and touch-screens can be individually tailored, or may mirror the main output.

Sound output will be rendered based on the interaction elements coming in from the control channels. Two types of sound will be employed: pre-composed musical tracks, and live environmental recordings. Each mobile user may have selected a music track as their "soundtrack" for their walk. Meanwhile, the microphone on the mobile phone serves to provide live sound capture of the outdoor environment. These elements will be correlated to similar elements from the other users, as well as history of previous users experiences to create a collage mix. Depending on the state of the user's sensors or location, as well as those of the others, the final sound output will be an evolving mix of deconstructed musical source elements injected with environmental noises, actively panning and substituting elements from other users and other moments while preserving a coherent sonic flow.

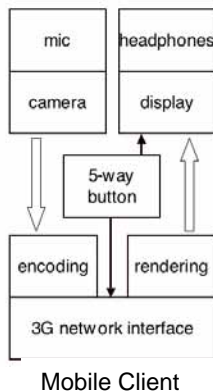
**ARCHITECTURE**

The technical infrastructure of Bastille Dérive will support both classical client-server architectures as well as peer-to-peer like end-to-end communication. The project server will be situated in the gallery space. There are three kinds of server side functions: audio/video (a/v) input/output (i/o), message services, and location services. There are two types of clients: the mobile devices and the small touch-screens in the gallery.

*Clients*

*Mobile Terminal*

The client will be a mobile telephone with multimedia capabilities. It has a network connection over 3G (or UMTS) packet switching mobile data infrastructures with a bandwidth throughput of up to 256 kbps. Localization services allow the The terminal is able to operate as a audio and video capture device, streaming live audio-video streams up the wireless network to the server in the gallery. It is also a sound/image display device, with MP3 audio decoding played over standard Walkman headphones, and with a 240x320 color graphics display. The client is able to read audio and video data streams from the network to display in real time. A custom Java program was developed, and has been downloaded to the phone's memory prior to the exhibition.



The user will interact with the system via a 5-way joystick-button system and two programmable soft-buttons. The phone folds switchblade style, concealing the numeric telephone keypad, but leaving the camera, display, audio jacks and interface buttons apparent. This configuration allows the phone to lose its telephone associations and be used as a nomadic a/v device.

While the primary mode of communication for the mobile client will be with the project server, there is the possibility of direct peer communication amongst the mobile devices. A profiling system will be used to construct a social networking map when each mobile device joins the system.

*In-gallery Touch Screen*

The stationary clients will be 12.5cm diagonal 768x1024 pixel handheld computers. They will be connected via 802.11g (WiFi) wireless networking to the server. One touch-screen interface corresponds to each mobile phone out in the field. The users will be able to manipulate the graphic user interface (GUI) on the touch-screen with their fingers. These in-gallery client devices serve as mediators between the mobile terminals and the final gallery audio/video projections.

## Servers

### A/V Server

The server will be programmed in MaxMSP/Jitter with several ancillary components. It is able to receive parallel audio/video upstreams from each of the three mobile phones. The server will have three distinct a/v output destinations: the mobile clients, the gallery clients, and the main projection display.

Audio to the mobile clients will be sent up to an Icecast server [16] running on the same machine that relays the audio stream to the mobile phones as an MP3 live stream. The incoming video from the mobile phones will be mirrored to each corresponding in-gallery client over the LAN via HTTP. The video stream is consolidated and encoded as an H.263 streamed over HTTP to the mobile clients. The monitor output as well as audio output of the server machine are connected to a local projection system in the gallery. This is the final combined output of the system, sent to a video projector and sound system.

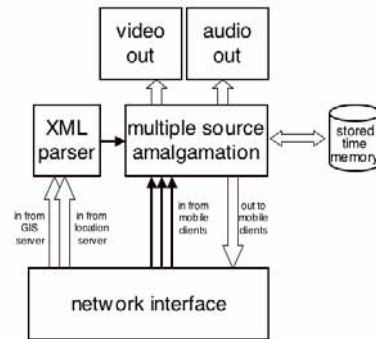


Figure. Server Architecture

### Messaging Server

The server will also receive control input resulting from the mobile user's actions on the telephone buttons, as well as the gallery participants' activities on the touch screens. These control messages will enter as OSC messages translated to XML over UDP and TCP socket connections.

### Location Server

Location based services are provided as a middleware layer server side between the mobile phone operator and the gallery. Each mobile client can be tracked geographically to a precision determined by the distance between two cell antennas. This location data is aggregated at the gallery server as one of the visual/musical parameters.

### Social Network Server

A user profile system allows for recording of user preferences and user-specific metadata. While on commercial systems, this kind of data would typically be used for personalization to feed recommendation engines, in the case of the our current framework, the data will be used as a means through which different users may enter into direct contact. This allows bootstrapping of a social network overlay to the system where principles from social computing are used to make connections between direct acquaintances, friends of friends, and users of similar profiles. This provides the peer-to-peer communications channel that runs parallel to the client-server media architecture predominant in the framework.